

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2000-099097

(43)Date of publication of application : 07.04.2000

(51)Int.Cl.

G10L 21/04

(21)Application number : 10-270244

(71)Applicant : SONY CORP

(22)Date of filing : 24.09.1998

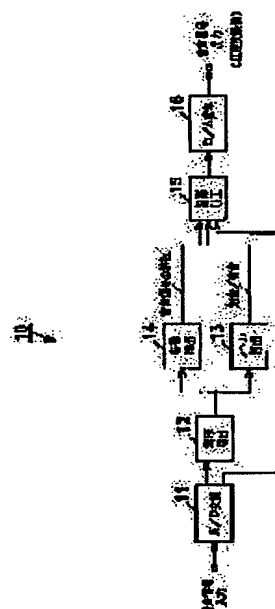
(72)Inventor : MURABAYASHI NOBORU
TAKAHASHI TAKAO

(54) SIGNAL REPRODUCING DEVICE AND METHOD, VOICE SIGNAL REPRODUCING DEVICE, AND SPEED CONVERSION METHOD FOR VOICE SIGNAL

(57)Abstract:

PROBLEM TO BE SOLVED: To enable converting reproducing speed of a voice signal without the degradation of the extent of understanding contents.

SOLUTION: A voice signal reproducing device 10 is provided with a waveform cutting out circuit 12 cutting out a signal waveform of a voice signal for each prescribed time unit, a level detecting circuit 13 detecting a signal level of a cut out signal waveform, a feature extracting circuit 14 extracting the feature of a cut out signal waveform, and a waveform processing circuit 15 processing a waveform of the voice signal by performing the deletion and/or addition of a signal waveform for each prescribed time unit based on a signal level of a signal waveform detected by the level detecting circuit 13 and features of a signal waveform extracted by the feature extracting circuit 14. In this voice signal reproducing device 10, reproducing speed of a voice signal is converted without changing pitch by deleting a part in which a soundless part and feature are continued.



(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号
特開2000-99097
(P2000-99097A)

(43) 公開日 平成12年4月7日 (2000. 4. 7)

(51) Int.Cl.⁷
G 1 0 L 21/04

識別記号

F I
G 1 0 L 3/02

テーマコード* (参考)
C 5 D 0 4 5

審査請求 未請求 請求項の数30 O L (全 14 頁)

(21) 出願番号 特願平10-270244

(22) 出願日 平成10年9月24日 (1998. 9. 24)

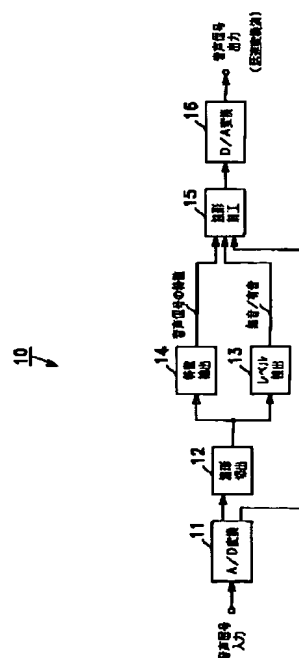
(71) 出願人 000002185
ソニー株式会社
東京都品川区北品川6丁目7番35号
(72) 発明者 村林 昇
東京都品川区北品川6丁目7番35号 ソニ
ー株式会社内
(72) 発明者 高橋 孝夫
東京都品川区北品川6丁目7番35号 ソニ
ー株式会社内
(74) 代理人 100067736
弁理士 小池 晃 (外2名)
Fターム(参考) 5D045 BA02 BB10

(54) 【発明の名称】 信号再生装置及び方法、音声信号再生装置、並びに、音声信号の速度変換方法

(57) 【要約】

【課題】 内容理解度の低下が無く音声信号の再生速度を変換できる音声信号再生装置を提供する。

【解決手段】 音声信号再生装置10は、音声信号の信号波形を所定の時間単位毎に切り出す波形切出回路12と、切り出した信号波形の信号レベルを検出するレベル検出回路13と、切り出した信号波形の特徴を抽出する特徴抽出回路14と、レベル検出回路13により検出した信号波形の信号レベルと特徴抽出回路14により抽出した信号波形の特徴とに基づき、所定の時間単位毎の信号波形の削除及び／又は追加を行って上記音声信号の波形加工をする波形加工回路15とを備える。この音声信号再生装置10では、無音部分と特徴が連続した部分を削除して、ピッチを変えずに音声信号の再生速度を変換する。



【特許請求の範囲】

【請求項1】 記録媒体から信号を再生する再生手段と、
上記再生手段により再生した再生信号の信号波形を所定の時間単位毎に切り出す信号切出手段と、
上記信号切出手段により切り出した信号波形の信号レベルを検出するレベル検出手段と、
上記信号切出手段により切り出した信号波形の特徴を抽出する特徴抽出手段と、
上記レベル検出手段により検出した所定の時間単位の信号波形の信号レベルと上記特徴抽出手段により抽出した所定の時間単位の信号波形の特徴とに基づき、所定の時間単位毎の信号波形の削除及び／又は追加を行って上記再生信号の波形加工をし、再生信号の時間軸を変換する時間軸変換手段とを備える信号再生装置。

【請求項2】 上記レベル検出手段により検出した所定の時間単位の信号波形の信号レベルと上記特徴抽出手段により抽出した所定の時間単位の信号波形の特徴とに応じて、上記記録媒体の再生制御をする制御手段を備えることを特徴とする請求項1に記載の信号再生装置。

【請求項3】 上記時間軸変換手段は、削除した信号波形の前後の信号波形を、重み付け関数を用いて波形加工して接続することを特徴とする請求項1に記載の信号再生装置。

【請求項4】 上記時間軸変換手段は、削除した信号波形の前後の信号波形を、その信号波形の微分値に基づき接続することを特徴とする請求項1に記載の信号再生装置。

【請求項5】 上記レベル検出手段は、所定の時間単位毎の信号波形の平均パワー及び／又は平均レベルから信号レベルを検出し、
上記時間軸変換手段は、信号レベルが所定の閾値以下である信号波形を、所定の時間単位毎に削除及び／又は追加することを特徴とする請求項1に記載の信号再生装置。

【請求項6】 上記特徴抽出手段は、所定の時間単位毎の信号波形の波形相関性からその信号波形の特徴を抽出し、
上記時間軸変換手段は、上記特徴が類似した信号波形を、所定の時間単位毎に削除及び／又は追加することを特徴とする請求項1に記載の信号再生装置。

【請求項7】 上記特徴抽出手段は、所定の時間単位毎の信号波形を周波数解析し、ピーク周波数の持続性からその信号波形の特徴を抽出し、
上記時間軸変換手段は、上記特徴が類似した信号波形を、所定の時間単位毎に削除及び／又は追加することを特徴とする請求項1に記載の信号再生装置。

【請求項8】 上記特徴抽出手段は、所定の時間単位毎の信号波形のレベル変化からその信号波形の特徴を抽出し、

上記時間軸変換手段は、上記特徴が類似した信号波形を、所定の時間単位毎に削除及び／又は追加することを特徴とする請求項1に記載の信号再生装置。

【請求項9】 記録媒体から信号を再生し、再生した再生信号の信号波形を所定の時間単位毎に切り出し、
切り出した信号波形の信号レベルを検出し、
切り出した信号波形の特徴を抽出し、
検出した所定の時間単位の信号波形の信号レベルと抽出した所定の時間単位の信号波形の特徴とに基づき、所定の時間単位毎の信号波形の削除及び／又は追加を行って上記再生信号の波形加工をし、再生信号の時間軸を変換することを特徴とする信号再生方法。

【請求項10】 検出した所定の時間単位の信号波形の信号レベルと抽出した所定の時間単位の信号波形の特徴とに応じて、上記記録媒体の再生制御をすることを特徴とする請求項9に記載の信号再生方法。

【請求項11】 削除した信号波形の前後の信号波形を、重み付け関数を用いて波形加工して接続し、再生信号の時間軸を変換することを特徴とする請求項9に記載の信号再生方法。

【請求項12】 削除した信号波形の前後の信号波形を、その信号波形の微分値に基づき接続し、再生信号の時間軸を変換することを特徴とする請求項9に記載の信号再生方法。

【請求項13】 所定の時間単位毎の信号波形の平均パワー及び／又は平均レベルから信号レベルを検出し、
上記信号レベルが所定の閾値以下である信号波形を、所定の時間単位毎に削除及び／又は追加することを特徴とする請求項9に記載の信号再生方法。

【請求項14】 所定の時間単位毎の信号波形の波形相関性からその信号波形の特徴を抽出し、
上記特徴が類似した信号波形を、所定の時間単位毎に削除及び／又は追加することを特徴とする請求項9に記載の信号再生方法。

【請求項15】 所定の時間単位毎の信号波形を周波数解析し、ピーク周波数の持続性からその信号波形の特徴を抽出し、
上記特徴が類似した信号波形を、所定の時間単位毎に削除及び／又は追加することを特徴とする請求項9に記載の信号再生方法。

【請求項16】 所定の時間単位毎の信号波形のレベル変化からその信号波形の特徴を抽出し、
上記特徴が類似した信号波形を、所定の時間単位毎に削除及び／又は追加することを特徴とする請求項9に記載の信号再生方法。

【請求項17】 音声信号の信号波形を所定の時間単位毎に切り出す信号切出手段と、
上記信号切出手段により切り出した信号波形の信号レベルを検出するレベル検出手段と、上記信号切出手段によ

り切り出した信号波形の特徴を抽出する特徴抽出手段と、

上記レベル検出手段により検出した所定の時間単位の信号波形の信号レベルと上記特徴抽出手段により抽出した所定の時間単位の信号波形の特徴とに基づき、所定の時間単位毎の信号波形の削除及び／又は追加を行って上記音声信号の波形加工をし、音声信号の再生速度を変換する速度変換手段とを備える音声信号再生装置。

【請求項18】 上記速度変換手段は、削除した信号波形の前後の信号波形を、重み付け関数を用いて波形加工して接続することを特徴とする請求項17に記載の音声信号再生装置。

【請求項19】 上記速度変換手段は、削除した信号波形の前後の信号波形を、その信号波形の微分値に基づき接続することを特徴とする請求項17に記載の音声信号再生装置。

【請求項20】 上記レベル検出手段は、所定の時間単位毎の信号波形の平均パワー及び／又は平均レベルから信号レベルを検出し、

上記速度変換手段は、信号レベルが所定の閾値以下である信号波形を、所定の時間単位毎に削除及び／又は追加することを特徴とする請求項17に記載の音声信号再生装置。

【請求項21】 上記特徴抽出手段は、所定の時間単位毎の信号波形の波形相関性からその信号波形の特徴を抽出し、

上記速度変換手段は、上記特徴が類似した信号波形を、所定の時間単位毎に削除及び／又は追加することを特徴とする請求項17に記載の音声信号再生装置。

【請求項22】 上記特徴抽出手段は、所定の時間単位毎の信号波形を周波数解析し、ピーク周波数の持続性からその信号波形の特徴を抽出し、

上記速度変換手段は、上記特徴が類似した信号波形を、所定の時間単位毎に削除及び／又は追加することを特徴とする請求項17に記載の音声信号再生装置。

【請求項23】 上記特徴抽出手段は、所定の時間単位毎の信号波形のレベル変化からその信号波形の特徴を抽出し、

上記速度変換手段は、上記特徴が類似した信号波形を、所定の時間単位毎に削除及び／又は追加することを特徴とする請求項17に記載の音声信号再生装置。

【請求項24】 音声信号の信号波形を所定の時間単位毎に切り出し、

切り出した信号波形の信号レベルを検出し、

切り出した信号波形の特徴を抽出し、

抽出した所定の時間単位の信号波形の信号レベルと抽出した所定の時間単位の信号波形の特徴とに基づき、所定の時間単位毎の信号波形の削除及び／又は追加を行って上記音声信号の波形加工をし、音声信号の再生速度を変換することを特徴とする音声信号の速度変換方法。

【請求項25】 削除した信号波形の前後の信号波形を、重み付け関数を用いて波形加工して接続し、音声信号の再生速度を変換することを特徴とする請求項24に記載の音声信号の速度変換方法。

【請求項26】 削除した信号波形の前後の信号波形を、その信号波形の微分値に基づき接続し、音声信号の再生速度を変換することを特徴とする請求項24に記載の音声信号の速度変換方法。

【請求項27】 所定の時間単位毎の信号波形の平均パワー及び／又は平均レベルから信号レベルを検出し、上記信号レベルが所定の閾値以下である信号波形を、所定の時間単位毎に削除及び／又は追加することを特徴とする請求項24に記載の音声信号の速度変換方法。

【請求項28】 所定の時間単位毎の信号波形の波形相関性からその信号波形の特徴を抽出し、

上記特徴が類似した信号波形を、所定の時間単位毎に削除及び／又は追加することを特徴とする請求項24に記載の音声信号の速度変換方法。

【請求項29】 所定の時間単位毎の信号波形を周波数解析し、ピーク周波数の持続性からその信号波形の特徴を抽出し、

上記特徴が類似した信号波形を、所定の時間単位毎に削除及び／又は追加することを特徴とする請求項24に記載の音声信号の速度変換方法。

【請求項30】 所定の時間単位毎の信号波形のレベル変化からその信号波形の特徴を抽出し、

上記特徴が類似した信号波形を、所定の時間単位毎に削除及び／又は追加することを特徴とする請求項24に記載の音声信号の速度変換方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、記録媒体に記録された信号を再生してその信号の時間軸を変換して出力する信号再生装置及び方法、並びに、音声信号の再生速度を通常の再生速度より早くしたり遅くしたりする特殊再生を行う音声信号再生装置及び音声信号の速度変換方法に関するものである。

【0002】

【従来の技術】一般に、家庭用ビデオテープレコーダ（VTR）には、ビデオテープを通常の再生速度よりも高速に再生し、短い時間で記録された映像や音声視聴することができる早送り再生機能が設けられている。しかしながら、この早送り再生機能を用いてビデオテープを再生した場合、出力される音声信号のピッチが変わってしまい、その音声を聞いても内容を理解することができなかった。そこで、近年のVTRでは、早送り再生時において音声信号の無音区間を削除する話速変換処理を行い、出力される音声信号のピッチを通常再生時のピッチと同一にし、その音声を聞いて内容を理解することができるようにしている。

【0003】

【発明が解決しようとする課題】ところが、無音区間を削除して話速変換処理をするVTRでは、より高速に早送り再生をした場合、無音区間が比較的多く含まれる音声に対しては有効に機能したが、無音区間があまり多くない音声に対しては有音区間も削除しなければならなく、出力された音声を聞いても内容を理解することができなかった。

【0004】本発明は、このような実情を鑑みてなされたものであり、信号の有効部分を削除することなく、時間軸を変換することができる信号再生装置及び方法を提供することを目的とする。

【0005】また、本発明は、内容理解度の低下が無く音声信号の再生速度を変換することができる音声信号再生装置及び音声信号の速度変換方法を提供することを目的とする。

【0006】

【課題を解決するための手段】本発明に係る信号再生装置は、記録媒体から信号を再生する再生手段と、上記再生手段により再生した再生信号の信号波形を所定の時間単位毎に切り出す信号切出手段と、上記信号切出手段により切り出した信号波形の信号レベルを検出するレベル検出手段と、上記信号切出手段により切り出した信号波形の特徴を抽出する特徴抽出手段と、上記レベル検出手段により検出した所定の時間単位の信号波形の信号レベルと上記特徴抽出手段により抽出した所定の時間単位の信号波形の特徴とに基づき、所定の時間単位毎の信号波形の削除及び／又は追加を行って上記再生信号の波形加工をし、再生信号の時間軸を変換する時間軸変換手段とを備えることを特徴とする。

【0007】この信号再生装置では、所定の時間単位毎に切り出した再生信号の信号波形から信号レベルと特徴とを検出し、この所定の時間単位毎に信号波形の削除及び／又は追加を行って上記再生信号の波形加工をし、再生信号の時間軸を変換する。

【0008】本発明に係る信号再生方法は、記録媒体から信号を再生し、再生した再生信号の信号波形を所定の時間単位毎に切り出し、切り出した信号波形の信号レベルを検出し、切り出した信号波形の特徴を抽出し、検出した所定の時間単位の信号波形の信号レベルと抽出した所定の時間単位の信号波形の特徴とに基づき、所定の時間単位毎の信号波形の削除及び／又は追加を行って上記再生信号の波形加工をし、再生信号の時間軸を変換することを特徴とする。

【0009】この信号再生方法では、所定の時間単位毎に切り出した再生信号の信号波形から信号レベルと特徴とを検出し、この所定の時間単位毎に信号波形の削除及び／又は追加を行って上記再生信号の波形加工をし、再生信号の時間軸を変換する。

【0010】本発明に係る音声信号の再生装置では、音

声信号の信号波形を所定の時間単位毎に切り出す信号切出手段と、上記信号切出手段により切り出した信号波形の信号レベルを検出するレベル検出手段と、上記信号切出手段により切り出した信号波形の特徴を抽出する特徴抽出手段と、上記レベル検出手段により検出した所定の時間単位の信号波形の信号レベルと上記特徴抽出手段により抽出した所定の時間単位の信号波形の特徴とに基づき、所定の時間単位毎の信号波形の削除及び／又は追加を行って上記音声信号の波形加工をし、音声信号の再生速度を変換する速度変換手段とを備えることを特徴とする。

【0011】この音声信号の再生装置では、所定の時間単位毎に切り出した音声信号の信号波形から信号レベルと特徴とを検出し、この所定の時間単位毎に信号波形の削除及び／又は追加を行って上記音声信号の波形加工をし、音声信号の再生速度を変換する。

【0012】本発明に係る音声信号の速度変換方法では、音声信号の信号波形を所定の時間単位毎に切り出し、切り出した信号波形の信号レベルを検出し、切り出した信号波形の特徴を抽出し、抽出した所定の時間単位の信号波形の信号レベルと抽出した所定の時間単位の信号波形の特徴とに基づき、所定の時間単位毎の信号波形の削除及び／又は追加を行って上記音声信号の波形加工をし、音声信号の再生速度を変換することを特徴とする。

【0013】この音声信号の速度変換方法では、所定の時間単位毎に切り出した音声信号の信号波形から信号レベルと特徴とを検出し、この所定の時間単位毎に信号波形の削除及び／又は追加を行って上記音声信号の波形加工をし、音声信号の再生速度を変換する。

【0014】

【発明の実施の形態】まず、本発明を適用した第1の実施の形態の話速変換装置について説明する。

【0015】図1に、本発明を適用した第1の実施の形態の話速変換装置のブロック構成図を示す。

【0016】図1に示す話速変換装置10は、例えば、ビデオテープレコーダ等の音声出力段に用いられ、早送り再生等がされたときに、音声信号の話速変換を行う装置である。

【0017】話速変換装置10は、アナログ／デジタル(A/D)変換回路11と、波形切出回路12と、レベル検出回路13と、特徴抽出回路14と、波形加工回路15と、デジタル／アナログ(D/A)変換回路16とを有している。

【0018】この話速変換装置10には、例えば、ビデオテープから再生されたアナログの音声信号が供給される。

【0019】A/D変換回路11は、入力されたアナログの音声信号をA/D変換してデジタルの音声信号に変換する。

【0020】波形切出回路12には、A/D変換回路11によりデジタルデータにされた音声信号が供給される。波形切出回路12は、時間的に連続している音声信号を所定の時間単位毎に分割して、その時間単位内の音声信号の信号波形を切り出す。信号波形を切り出す時間単位は、例えば、10msec～200msec程度である。この時間単位のことを以後音声ブロックと呼ぶ。

【0021】レベル検出回路13には、波形切出回路12により信号波形が切り出された音声信号、音声ブロック毎に供給される。レベル検出回路13は、音声ブロック毎に信号信号のレベルを検出して、その音声ブロックが無音区間であるか有音区間であるかを判別する。例えば、レベル検出回路13は、この音声ブロック毎の音声信号の平均パワー（電力）P或いは平均レベルMを求め、この平均パワーP或いは平均レベルMが所定の閾値より高ければ有音区間と判断し、所定の閾値より低ければ無音区間と判断する。音声ブロック内の音声信号の平均パワーP及び平均レベルMは、以下のように算出することができる。

$$\begin{aligned} \text{【0022】平均パワー } P &= (1/N) \sum i^2 \\ \text{平均レベル } M &= (1/N) \sum |i| \end{aligned}$$

但し、累積加算は音声ブロック内で行い、Nは音声ブロック内のサンプリングデータ数、iは音声信号の信号レベル（振幅）である。

【0023】特徴抽出回路14には、波形切出回路12により信号波形が切り出された音声信号が、音声ブロック毎に供給される。特徴抽出回路14は、供給された音声信号の信号波形から、音声ブロック毎に音声信号の特徴を抽出する。音声信号の特徴とは、例えば、音声信号のピッチや音声信号の周波数特性等である。

【0024】波形加工回路15には、波形切出回路12により信号波形が切り出された音声信号が、音声ブロック毎に供給される。また、この波形加工回路15には、レベル検出回路13より判断された音声ブロックが有音区間であるか無音区間であるかの判断結果、及び、特徴抽出回路14により抽出された音声ブロック毎の音声信号の特徴が供給される。

【0025】波形加工回路15は、音声ブロックが有音区間であるか無音区間であるかの判断結果と音声信号の特徴とに基づき波形加工処理を行い、音声信号の再生速度変換を行う。

【0026】D/A変換回路16には、波形加工回路15により波形加工がされたデジタルの音声信号が供給される。D/A変換回路16は、デジタルの音声信号をアナログ信号に変換して出力する。

【0027】以上のような構成の話速変換装置10では、ビデオテープの早送り再生をした場合、波形加工回路15が音声信号の加工処理をすることにより、音声のピッチを変化させず音声信号の話速度を速くし、早聞きを可能としている。

【0028】波形加工回路15の処理内容をさらに詳細に説明する。

【0029】波形加工回路15は、無音区間の音声ブロックを削除する。さらに、波形加工回路15は、特徴が類似した音声信号が連続している音声ブロックの一部を削除する。このとき、波形加工回路15は、音声信号の再生速度に応じて音声ブロックの削除の割合を決定して、削除をしていく。すなわち、波形加工回路15は、早送り再生をする際の速度に応じて削除するブロック数を決定する。そして、波形加工回路15は、削除していない残っている音声ブロック同士を接続して出力し、音声信号の再生速度の変換処理すなわち話速度の変換処理を行う。

【0030】ここで、波形加工回路15は、削除していない残っている音声ブロック同士を接続する際にそのまま単純に接続しては聴覚上問題が生じるので、以下のように滑らかに接続をする。

【0031】例えば、図2に示すように、時間的に連続した音声ブロックA、B、Cがあり、その真ん中の音声ブロックBを削除するものとする。音声ブロックAと音声ブロックCを単純に接続したのでは、音声ブロックAの信号波形の最後の部分（時刻t1における信号）と音声ブロックCの信号波形の最先の部分（時間t2における信号）とが不連続となり、この不連続部分が雑音になる場合がある。

【0032】そこで、波形加工回路15は、例えば、音声ブロックAの最先の部分（時刻ta）で1、音声ブロックBの最後の部分（時刻t2）で0となるような波形接続重み付け関数fa(t)を、音声ブロックAの音声信号に乗ずる。また、音声ブロックBの最先の部分（時刻t1）で0、音声ブロックCの最後の部分（時刻tc）で1となるような波形接続重み付け関数fc(t)を、音声ブロックCの音声信号に乗ずる。そして、この重み付け関数を乗じた音声ブロックAと音声部録Cの音声信号を接続する。

【0033】具体的には、音声ブロックAの音声信号に乗ずる重み付け関数fa(t)、及び、音声ブロックCの音声信号に乗ずる重み付け関数fc(t)は、以下のようなになる。

$$f_a(t) = -(t / (t_2 - t_a)) + t_2 / (t_2 - t_a)$$

$$f_c(t) = (t / (t_c - t_1)) - t_1 / (t_c - t_1)$$

そして、音声ブロックA～Bの音声信号をAB(t)、音声ブロックB～Cの音声信号をBC(t)とすると、波形接続後の信号AC(t)は、以下のようなになる。

$$AC(t) = AB(t) \cdot f_a(t) + BC(t) \cdot f_c(t)$$

波形加工回路15では、このように波形接続重み付け関数を用いて処理することにより、削除していない音声ブ

ロック同士を滑らかに接続でき、比較的聴覚上違和感のない音声信号接続が行える。

【0034】なお、波形接続重み付け関数 $f(t)$ として、 $f(t) = at + b$ のような線形1次関数を例に挙げたが、その他に $f(t) = at^2 + b$ のような線形2次関数、或いは、 $f(t) = a \cdot \exp(-t/\tau) + b$ のような指数関数であってもよい。

【0035】また、波形加工回路15では、図3に示すように、音声ブロックAの接続点となるa点を音声ブロックCの音声信号に接続する場合、このa点における微分係数の符号と異なる微分係数のc1点で接続するのではなく、このa点における微分係数の符号と同一の符号となるc2点で接続するように、音声ブロックBを削除するようにしてもよい。

【0036】つぎに、この話速変換装置10により実際の音声信号を話速変換した場合の信号波形について、図4～図7の波形図を用いて説明する。

【0037】図4は、入力音声信号の波形を示す図である。図5は、話速変換した後の出力音声信号の波形を示す図である。図6は、上記図4に示した入力音声信号の波形の一部分を拡大した図である。図7は、上記図5に示した出力音声信号の波形の一部分を拡大した図である。図4及び図5は、横軸の1目盛が5000サンプルを示しており、時間にして約0.104秒となっている。また、図6及び図7は、横軸の1目盛が1000サンプルを示しており、時間にして約0.0208秒となっている。また、音声信号は、サンプリング周波数 $f_s = 48000 \text{ KHz}$ 、量子化ビット数16ビットの信号である。

【0038】これら各図を見て分かるように、波形削除処理を行い話速処理を行って、波形接続処理を行っても、入力音声信号と出力音声信号のピッチが変化しておらず、また音声波形も滑らかに接続されている。

【0039】ここで、図8及び図9に示す波形のX部分は音声信号の子音区間であり、波形のY部分は母音区間である。入力音声信号と出力音声信号の子音区間を比べ*

*ると、話速変換処理を行った後の出力音声信号の方が子音区間が短くなっていることが分かる。同様に、入力音声信号と出力音声信号の母音区間を比べると、話速変換処理を行った後の出力音声信号の方が母音区間が短くなっていることが分かる。

【0040】そのため、話速処理を行って音声波形を削除する場合に、削除する割合をあまり多くしてしまうと会話音声の内容が分からなくなってしまう。

【0041】実験の結果によると、例えば母音区間については入力音声信号のピッチが約1.0λ程度あったとすれば、出力音声信号は、母音区間に1/2～1/3程度のピッチがあれば、会話内容の劣化がほとんどなく良好となる。また、子音区間についてもほぼ同様である。

【0042】以上のように本発明の第1の実施の形態の話速変換装置10では、音声ブロック毎に切り出した音声信号の信号波形から、音声信号の信号レベルと音声信号の特徴とを検出し、音声ブロックの削除を行って上記音声信号の波形加工をすることにより、内容理解度の低下が無く音声信号の再生速度を変換することができる。

【0043】なお、波形加工回路15において音声ブロックを削除する場合、例えば2倍速再生をするのであればブロック数を全体の1/2に間引けばよく、無音ブロックが非常に長いからといって必要以上のブロック数を削除することはない。また、図示しないビデオテープの回転速度のコントローラ等に音声信号の特徴等をフィードバックして、無音区間が多い部分では例えば3倍速や4倍速といった非常に早い早送り再生をし、有音区間では音声信号の内容が理解できる程度の早送り再生をするといったような可変速再生を行ってもよい。

【0044】また、波形加工回路15は、早送り再生の再生速度に応じて、音声ブロックの削除方法を変えても良い。例えば、下記の表に示すように、低速度（1～1.5倍速）、中速度（1.5～2.5倍速）、高速度（2.5倍以上）でそれぞれ削除方法を変えても良い。

【0045】

【表1】

再生速度	低速度 (1～1.5 倍速)	中速度 (1.5～2.5 倍速)	高速度 (2.5 倍速以上)
処理方法	無音削除	無音削除 及び 母音部削除	無音削除 母音部削除 子音部削除

【0046】なお、以上の処理で所定の速度にならない場合は、例えば、低速度であれば母音区間の削除を行

い、中速度であれば子音区間の削除を行い、また、高速度であれば音声レベルの低い部分を削除するようにしても良い。

【0047】また、高速度の高速再生の場合における母音区間と子音区間との検出については、まず母音区間の検出を行い、その前部の音声区間を子音区間とするなどの処理を行う。母音区間の検出については、自己相関関数によるピッチ検出を行い、その検出区間を母音区間とするなどの処理を行う。また、この場合、音声信号をそのまま自己相関処理する他に、対数スペクトル処理の後、自己相関処理を行い、ピッチ検出するようにしても良い。

【0048】つぎに、本発明を適用した第2の実施の形態の話速変換装置について説明する。

【0049】図8に、本発明を適用した第2の実施の形態の話速変換装置のブロック構成図を示す。なお、この第2の実施の形態の話速変換装置を説明するのにあたり、上記第1の実施の形態の話速変換装置10と同一の回路については、図面中に同一の符号を付け、その詳細な説明を省略する。また、第3の実施の形態以降も同様とする。

【0050】図8に示す話速変換装置20は、例えば、ビデオテープレコーダ等の音声出力段に用いられ、早送り再生等がされたときに、音声信号の話速変換を行う装置である。

【0051】話速変換装置20は、アナログ/デジタル(A/D)変換回路11と、波形切出回路12と、レベル検出回路13と、相関性検出回路21と、波形加工回路15と、デジタル/アナログ(D/A)変換回路16とを有している。波形加工回路15は、間引き回路22と、波形接続回路23とから構成されている。

【0052】相関性検出回路21には、波形切出回路12により信号波形が切り出された音声信号が、音声ブロック毎に供給される。相関性検出回路21は、波形切出回路12により切り出した音声ブロック間の自己相関関数を求める。

【0053】間引き回路22は、相関性検出回路21で求めた相関性のある音声ブロック、及び、レベル検出回路13で求めた無音区間の音声ブロックを削除する。

【0054】波形接続回路23は、削除せず残った音声ブロック同士を音声波形が滑らかに接続するように接続処理を行う。

【0055】本発明の第2の実施の形態の話速変換装置20では、このように波形の相関性を検出することによって、音声信号が類似した部分を検出することができ、検出した音声信号の類似部分を削除することによって、話速変換をすることができる。

【0056】つぎに、本発明を適用した第3の実施の形態の話速変換装置について説明する。

【0057】図9に、本発明を適用した第3の実施の形

態の話速変換装置のブロック構成図を示す。

【0058】図9に示す話速変換装置30は、例えば、ビデオテープレコーダ等の音声出力段に用いられ、早送り再生等がされたときに、音声信号の話速変換を行う装置である。

【0059】話速変換装置30は、アナログ/デジタル(A/D)変換回路11と、波形切出回路12と、レベル検出回路13と、レベル比較回路31と、波形加工回路15と、デジタル/アナログ(D/A)変換回路16とを有している。

【0060】レベル比較回路31には、波形切出回路12により信号波形が切り出された音声信号が、音声ブロック毎に供給される。レベル比較回路31は、音声ブロック毎に、音声信号のレベルを検出する。

【0061】間引き回路32は、無音区間の音声ブロックを削除するとともに、レベル比較回路31のレベル検出結果に基づき、音声ブロック毎にレベル検出を行い、レベルが同じ音声ブロックが所定回数続いたら、その音声ブロックのうちいくつかのブロックを削除する。削除するブロックの数は、再生速度等に応じて決定しても良く、また、連続した同レベルの音声ブロックのうち、レベルが異なる音声ブロックに隣接する2つの音声ブロックを残しその他の音声ブロックを削除しても良い。

【0062】波形接続回路23は、削除せず残った音声ブロック同士を音声波形が滑らかに接続するように接続処理を行う。

【0063】例えば、この話速変換装置30は、図10に示すように、A～Fの音声ブロックが入力され、音声ブロックB、C、Dのレベルが同一だった場合、音声ブロックCを削除し、BとD区間の間で波形接続処理を行う。また、音声ブロックB、C、D、Eのレベルが同一だった場合、音声ブロックCとDを削除し、BとEとの間で波形接続処理を行う。

【0064】本発明の第3の実施の形態の話速変換装置30では、このように音声レベルを検出することによって、音声信号が類似した部分を検出することができ、検出した音声信号の類似部分を削除することによって、話速変換をすることができる。

【0065】つぎに、本発明を適用した第4の実施の形態の話速変換装置について説明する。

【0066】図11に、本発明を適用した第4の実施の形態の話速変換装置のブロック構成図を示す。

【0067】図11に示す話速変換装置40は、例えば、ビデオテープレコーダ等の音声出力段に用いられ、早送り再生等がされたときに、音声信号の話速変換を行う装置である。

【0068】話速変換装置40は、アナログ/デジタル(A/D)変換回路11と、波形切出回路12と、レベル検出回路13と、周波数解析回路41と、ピーク周波数検出回路42と、ピーク周波数継続性検出回路43

10

20

30

40

50

と、波形加工回路15と、デジタル／アナログ（D／A）変換回路16とを有している。

【0069】周波数解析回路41には、波形切出回路12により信号波形が切り出された音声信号が、音声ブロック毎に供給される。周波数解析回路41は、音声ブロック毎に、音声信号の周波数解析を行う。

【0070】ピーク周波数検出回路42は、周波数解析回路41の解析結果に基づき、音声ブロック毎に、第1ピーク周波数と第2のピーク周波数とを検出する。

【0071】ピーク周波数継続性検出回路43は、ピーク周波数検出回路42の検出結果に基づき、ピーク周波数が同一の音声ブロックが継続しているかどうかを検出する。

【0072】間引き回路44は、無音区間の音声ブロックを削除するとともに、ピーク周波数継続性検出回路43の検出結果に基づき、ピーク周波数が同じ音声ブロックが所定回数続いたら、その音声ブロックのうちいくつかのブロックを削除する。削除するブロックの数は、再生速度等に応じて決定しても良く、また、連続した同じピーク周波数の音声ブロックのうち、ピーク周波数が異なる音声ブロックに隣接する2つの音声ブロックを残しその他の音声ブロックを削除しても良い。

【0073】波形接続回路23は、削除せず残った音声ブロック同士を音声波形が滑らかに接続するように接続処理を行う。

【0074】本発明の第4の実施の形態の話速変換装置40では、このようにピーク周波数を検出することによって、音声信号が類似した部分を検出することができ、検出した音声信号の類似部分を削除することによって、話速変換をすることができる。

【0075】つぎに、本発明を適用した第5の実施の形態について説明をする。この第5の実施の形態は、上述した本発明の第1から第4の実施の形態の話速変換装置を適用したディスク再生装置である。

【0076】図12に、本発明を適用した第5の実施の形態のディスク再生装置のブロック構成図を示す。

【0077】図12に示すディスク再生装置50は、例えば、映像及び音声デジタル記録された光ディスク等を再生し、ユーザにより早送り再生操作等がされた場合には、映像信号を高速再生するとともに、音声信号の話速変換を行うことができる装置である。

【0078】ディスク再生装置50は、光ディスク51に記録された信号を読み取りデジタルデータにする再生処理回路52と、再生処理回路52により読み取られたデジタルデータを映像信号と音声信号とに分離するデマルチプレクサ53とを有している。

【0079】また、ディスク再生装置50は、デマルチプレクサ53により分離された映像信号が供給され、映像信号のデコード処理やエラー訂正処理等を行う映像信号処理回路54と、早送り再生時等に映像信号の間引き

処理等を行う画像加工回路55と、画像加工回路44により加工された映像信号をアナログ信号に変換し出力する映像用D／A変換回路56とからなる映像再生系を有している。この映像再生系では、ユーザにより早送り再生操作がされると、画像加工回路55がフレームの間引き処理等を行って、所定の再生速度の映像信号を出力する。

【0080】また、ディスク再生装置50は、デマルチプレクサ53により分離された音声信号が供給され、音声信号のデコード処理やエラー訂正処理を行い、並びに、音声信号を所定の時間単位毎の音声ブロックで波形の切り出し処理を行う音声信号処理回路57と、レベル検出回路13と、特徴抽出回路14と、波形加工回路15と、音声用D／A変換回路16とを有している。

【0081】また、ディスク再生装置50は、各回路の制御及びユーザからの操作入力を受け付けるシステムコントローラ58と、システムコントローラ58の制御に基づき光ディスク51のサーボコントロールを行うサーボ制御回路59とを有している。

【0082】このような構成のディスク再生装置50では、ユーザにより早送り再生の操作処理がされると、画像加工回路55が画像データを加工して、早送り再生画像を出力する。また、波形加工回路14が、無音の音声ブロック及び特徴が連続した音声ブロックを削除して、早送り再生画像に対応した再生速度の音声信号を出力する。また、システムコントローラ58は、レベル検出回路13からの音声ブロックが無音であるか有音であるかの判断結果、及び、特徴抽出回路14の音声信号の特徴に基づき、サーボ制御回路59を介して光ディスク51の回転速度を制御する。このことにより、例えば、無音区間が多い部分では例えば3倍速や4倍速といった非常に早い早送り再生をし、有音区間では音声信号の内容が理解できる程度の早送り再生をするといったような可変速再生を行うことができる。

【0083】本発明の第5の実施の形態のディスク再生装置50では、このように光ディスク51に記録された音声信号に対しても、話速変換をすることができる。

【0084】以上、本発明の第1から第5の実施の形態について説明した。各実施の形態の話速変換装置並びにディスク再生装置では、無音部分の削除とともに、波形相関性、ピーク周波数特性、レベル変化等の音声信号の特徴を検出し、その特徴が連続した部分を削除している。このことにより、音声ピッチを通常時と変えることなく音声信号の再生速度を変更することができる。

【0085】また、各実施の形態では、音声信号の一部を削除することにより、高速再生を実現する例を示したが、例えば、無音部分を検出した場合にはその無音部分にさらに無音部分を追加したり、また、類似した特徴が連続した部分を検出した場合にはその特徴と類似した特徴の信号を追加したりすることにより、通常速度より遅

10

20

30

40

50

い低速再生を実現することができる。具体的には、波形加工回路14において、音声ブロックを追加し、追加した音声ブロックを接続処理することにより実現できる。

【0086】また、本実施の形態では、音声信号の速度変換をする例について説明したが、本発明は、音声信号に限らず、例えば画像信号等であってもよい。

【0087】

【発明の効果】本発明では、所定の時間単位毎に切り出した再生信号の信号波形から信号レベルと特徴とを検出し、この所定の時間単位毎に信号波形の削除及び／又は追加を行って上記再生信号の波形加工をし、再生信号の時間軸を変換する。このことにより本発明では、信号の有効部分を削除することなく、時間軸を変換することができる。

【0088】また、本発明では、所定の時間単位毎に切り出した音声信号の信号波形から信号レベルと特徴とを検出し、この所定の時間単位毎に信号波形の削除及び／又は追加を行って上記音声信号の波形加工をし、音声信号の再生速度を変換する。このことにより本発明では、簡易な構成で音声ピッチの変動を無くすことができ、内容理解度の低下が無く音声信号の再生速度を変換することができる。

【図面の簡単な説明】

【図1】本発明の第1の実施の形態の話速変換装置のブロック構成図である。

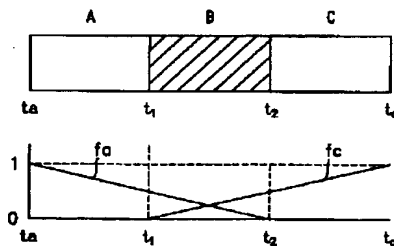
【図2】上記話速変換装置の波形加工回路による、重み付け関数を用いた音声信号の接続処理を説明するための図である。

10

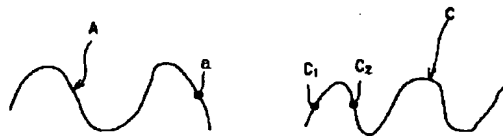
20

*

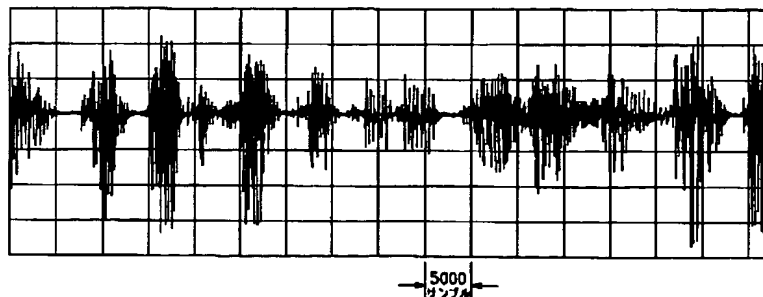
【図2】



【図3】



【図4】



入力音声信号

*【図3】上記話速変換装置の波形加工回路による、微分符号を用いた音声信号の接続処理を説明するための図である。

【図4】上記話速変換装置に入力される音声信号の一例を示す波形図である。

【図5】図4で示した音声信号を上記話速変換装置により話速変換を行った後の音声信号を示す波形図である。

【図6】図4で示した音声信号を拡大した図である。

【図7】図5で示した音声信号を拡大した図である。

【図8】本発明の第2の実施の形態の話速変換装置のブロック構成図である。

【図9】本発明の第3の実施の形態の話速変換装置のブロック構成図である。

【図10】上記第3の実施の形態の話速変換装置の波形加工回路による、音声信号の削除処理を説明するための図である。

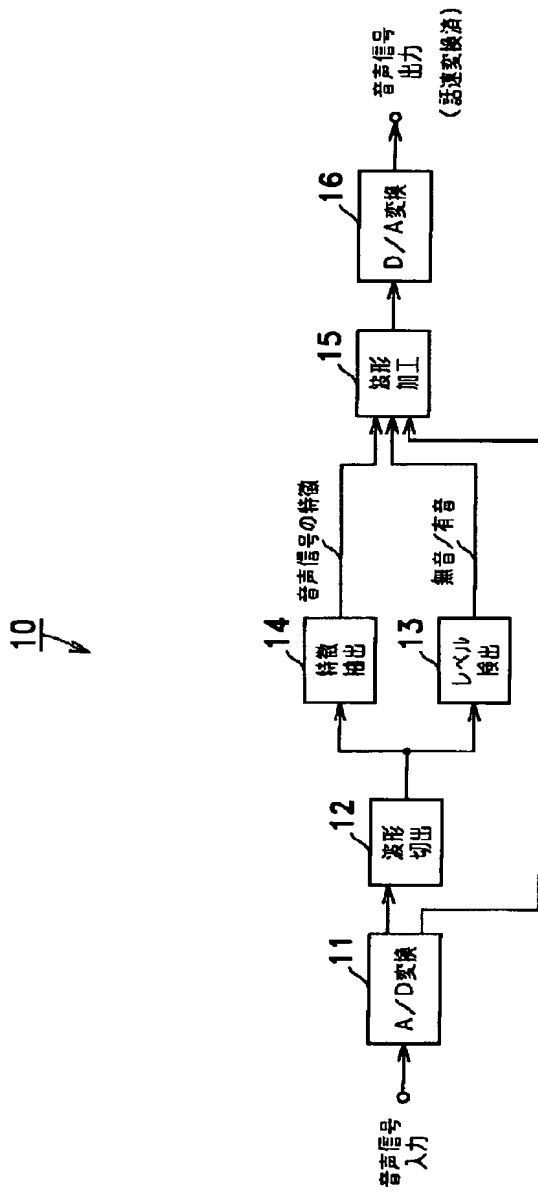
【図11】本発明の第4の実施の形態の話速変換装置のブロック構成図である。

【図12】本発明の第5の実施の形態のディスク再生装置のブロック構成図である。

【符号の説明】

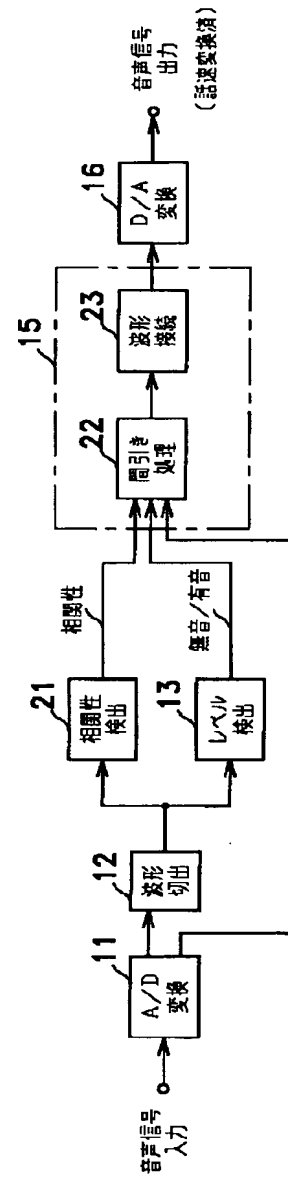
10, 20, 30, 40 話速変換装置、12 波形切出回路、13 レベル検出回路、14 特徴抽出回路、15 波形加工回路、21 相関性検出回路、31 レベル比較回路、41 周波数解析回路、42 ピーク周波数検出回路、43 ピーク周波数継続性検出回路、50 ディスク再生装置

【図1】

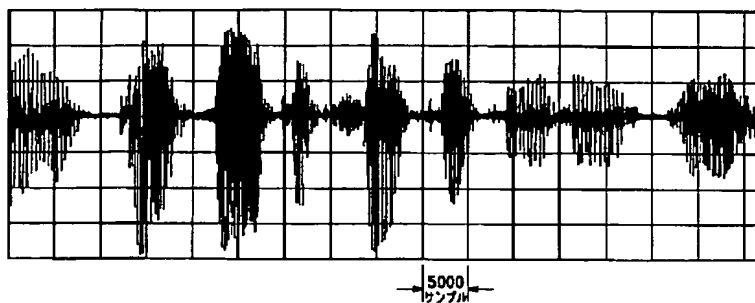


20

【図8】

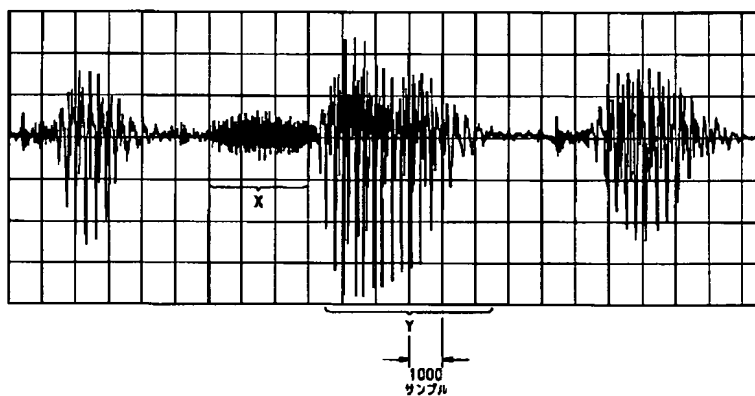


【図5】



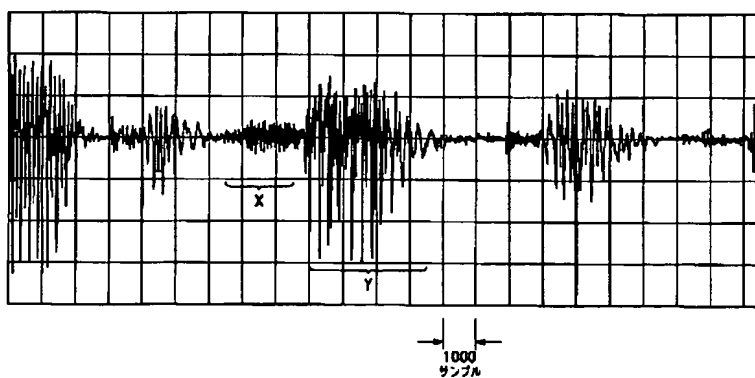
出力音声信号

【図6】



入力音声信号

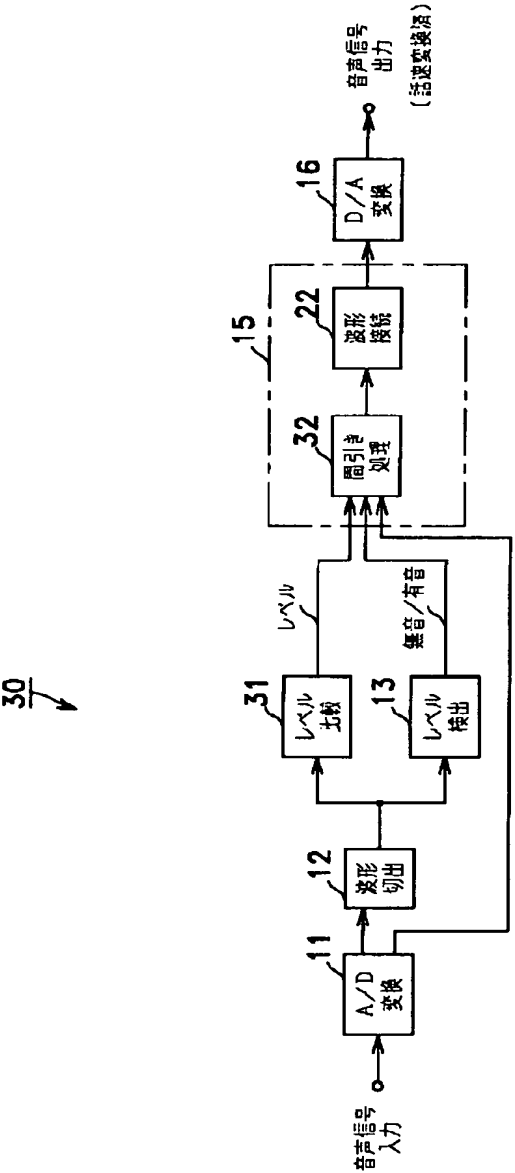
【図7】



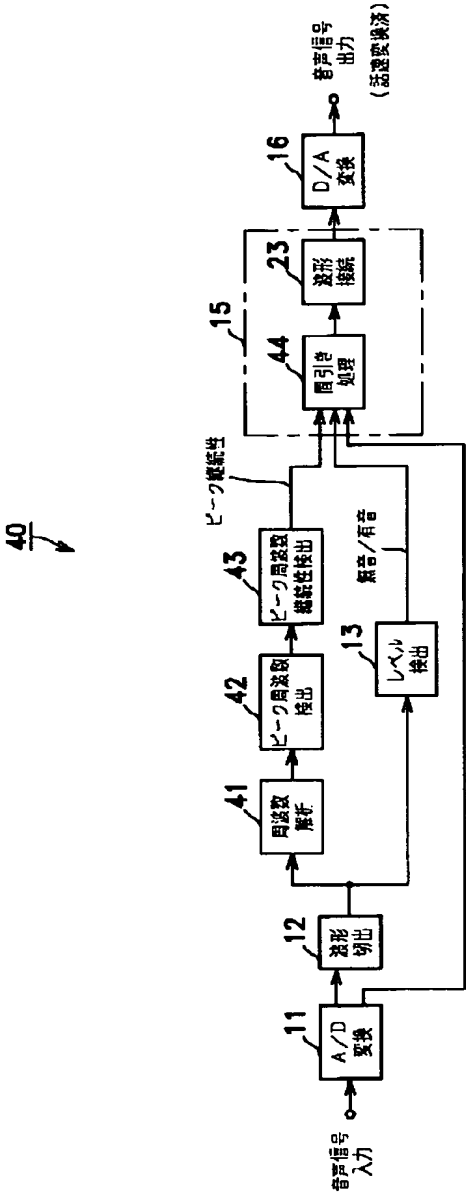
出力音声信号

(12)

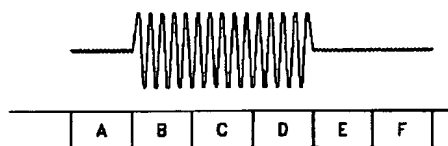
【図9】



【図11】



【図10】



【図12】

50

